

Data Science Academy – internal capacity development program

Dominika Rogalińska – Director of Innovation Department,
Statistics Poland - author

Anna Borowska – Civil Service Director,
Statistics Poland - presenter

HRMT workshop, 14 – 16 October 2024, Geneva, Switzerland

What was our inspiration?

UN-CEBD TT on Training, Competencies and Capacity Development

- **Maturity Matrix**
- **Competency Framework**
- **Training Catalogue**
- **Support in training development**
- **Work for Regional Hubs**

ONS-UNECE ML Group

- **International research collaboration**
- **Platform for knowledge sharing**
- **Focus on Machine Learning in official statistics**

ESSNet Trusted Smart Statistics - Web Intelligence Network

- **Building network of NSIs from ESS countries**
- **Focus on web data**
- **Emphasis on training and networking**

Data Science Academy @Statistics Poland (1)

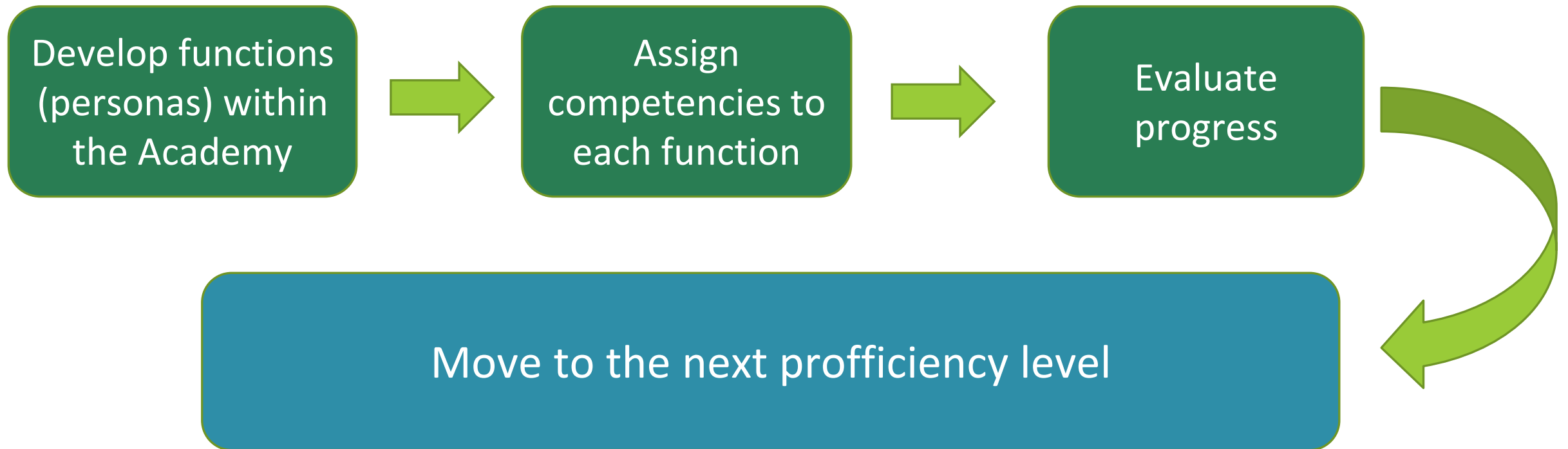
Internal development program

```
graph TD; A[Internal development program] --> B[Build a modern organization, able to anticipate and target fast-changing information needs by tapping into a multitude of data sources]; B --> C[Increase the integration of administrative and big data sources in statistical production – experimental & official];
```

Build a modern organization, able to anticipate and target fast-changing information needs by tapping into a multitude of data sources

Increase the integration of administrative and big data sources in statistical production – experimental & official

Data Science Academy @Statistics Poland (2)



Data Science Academy @Statistics Poland (3)

Data Scientist

Research design

Data acquisition

Data management

Data cleaning

Data integration

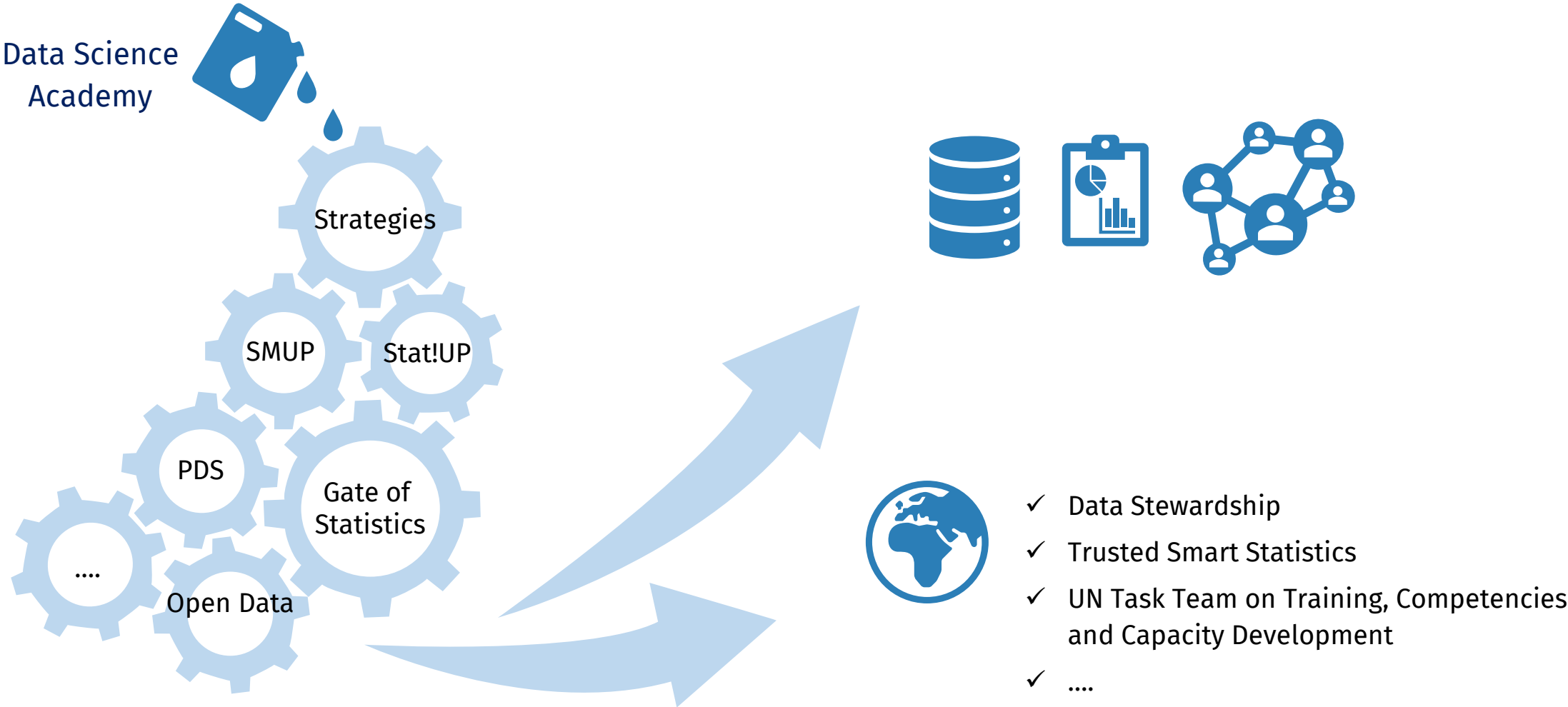
Data analyst

Data analysis

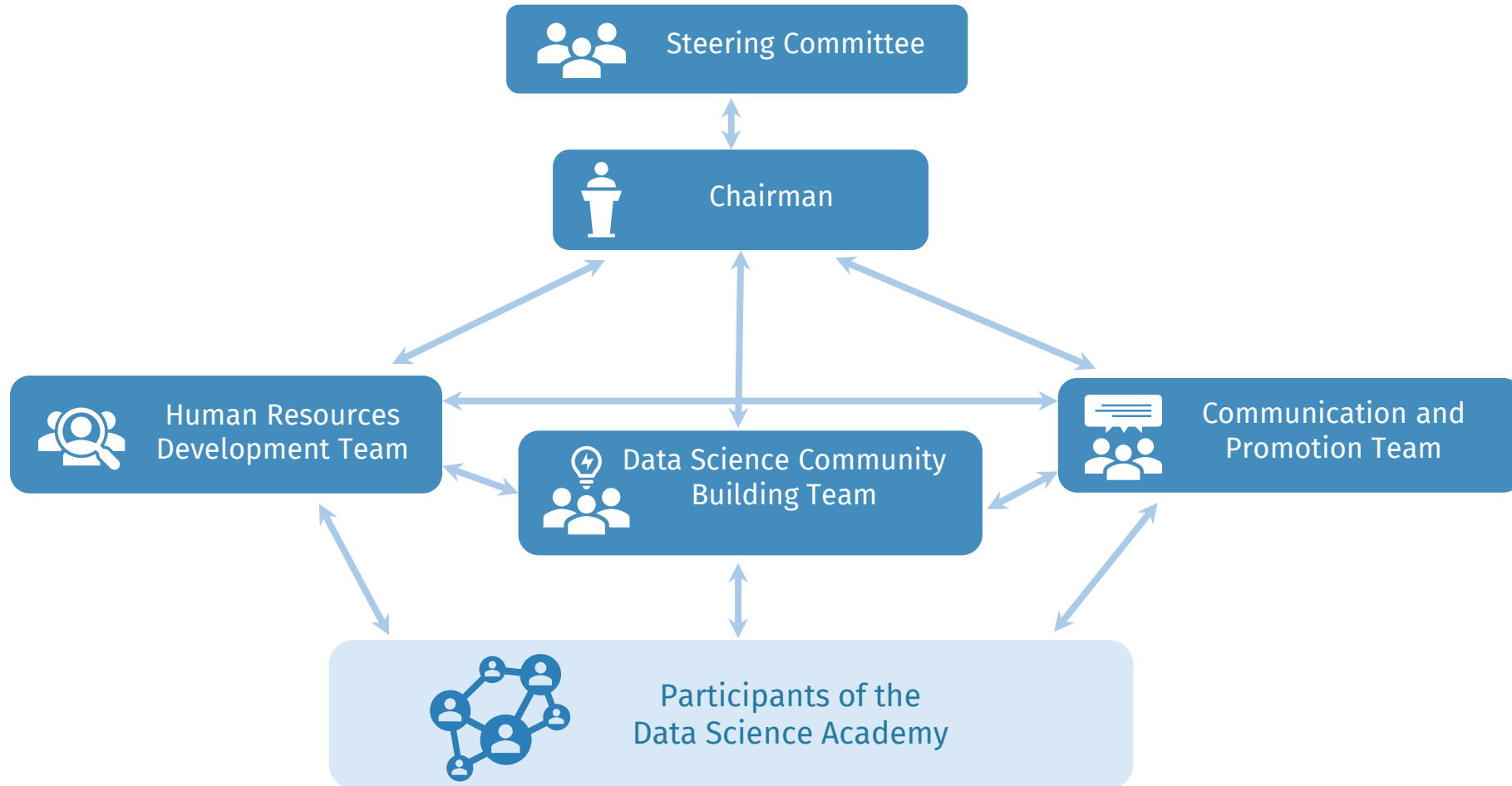
Data dissemination

Data engineer

Data Science Academy vs. strategic activities



Organizational aspects – A Steering Committee



Human Resources Development Team (1)



Human Resources Development Team (2)

01 - RECRUITMENT

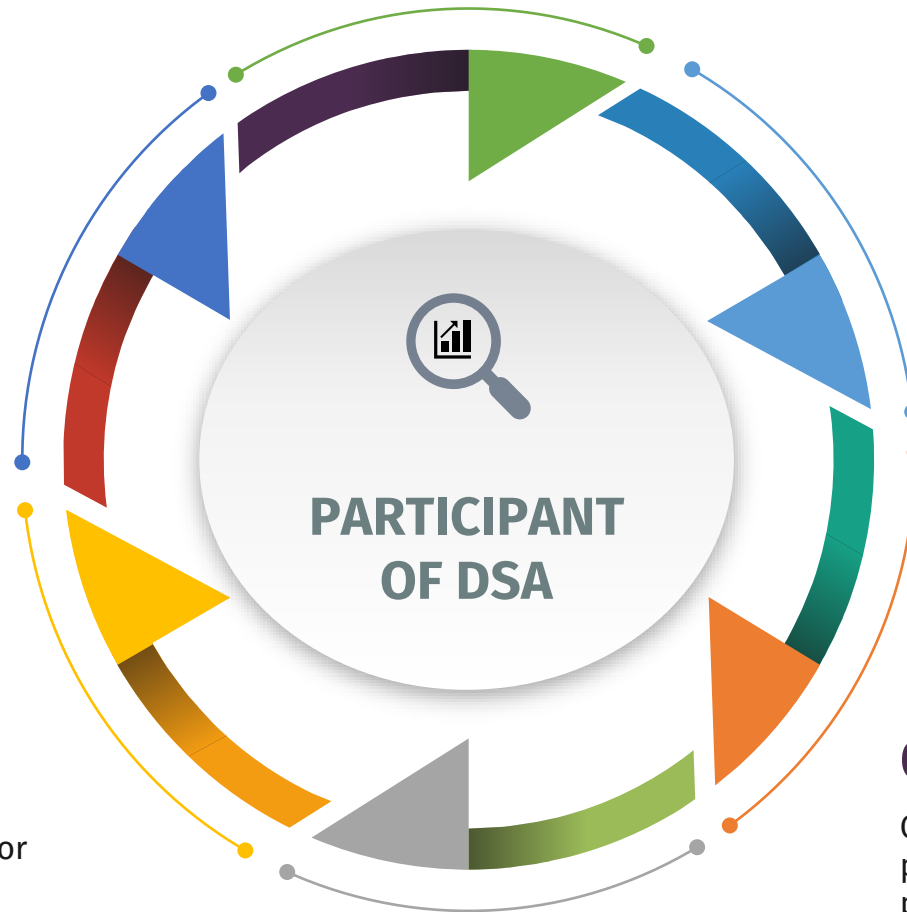
Verification of candidate applications; motivational meetings; English proficiency tests; content tests; recommendation; increased salary.

02 - ON-BOARDING

Organizational meetings for participants and superiors; implementation into processes implemented in DAS; presenting the principles of cooperation (building a culture of work organization).

03 - ASSIGNING TASKS

Possibility to submit a test/task or assign a test/task; registration of the project (research/task) in the GitLab environment; mandatory training with GitLab – preparation for working in the agile methodology; building effective teams; internal mobility of participants; reporting and monitoring work results.



04 – ASSESSMENT 180

Competence assessment according to the model adopted at the Data Science Academy; including elements of the participant's self-assessment and supervisor's assessment; feedback based on final calibration; two-year cycle.

05 – DEVELOPMENT CONTRACT

Development meetings with participants; preparation of individual development contracts; taking into account short-term and long-term goals; defining the expected effect after evaluation in ADS; two-year cycle.

06 - EVALUATION

Collecting all participant activities in one cycle; preparing recommendations based on the participant's needs, defined criteria and level of involvement in the Academy's activities.

Human Resources Development Team (3)

List of Data Science Academy participants

- Recommendation of candidates;
- Meetings with candidates; motivation research;
- Final approval of the list of participants;
- Announcement of results (information sent via e-office)

Recruitment for roles:
Data Scientist, Data Engineer, Data Analyst, Line Manager

Expert:

a person with at least 3 years of experience in public statistics, with extensive experience in the use of non-statistical data sources on the domestic and international market (at least one year of work, e.g. in international working groups), with high analytical, programming and interpersonal skills.

Specialist:

a person with at least one year of experience in public statistics, with experience in the use of non-statistical data sources on the domestic market with high analytical and/or programming skills.

based on management's recommendation (recommendation form)

examining motivation to enter the Academy, take up new challenges, increase commitment, share knowledge, develop in a selected role vs. supervisor's recommendation sheet

approval of the program management, the number of participants depends on the budget

Recommendation form

List of criteria

1. Indication by the superior/justification of the recommendation (role/level).
2. Appropriate work experience (min. 1 year) and commitment.
3. A high level of aspiration and readiness for the challenges that await the new role.
4. Currently achieved results.
5. Everyday interest in the field of data science.
6. Motivation research.

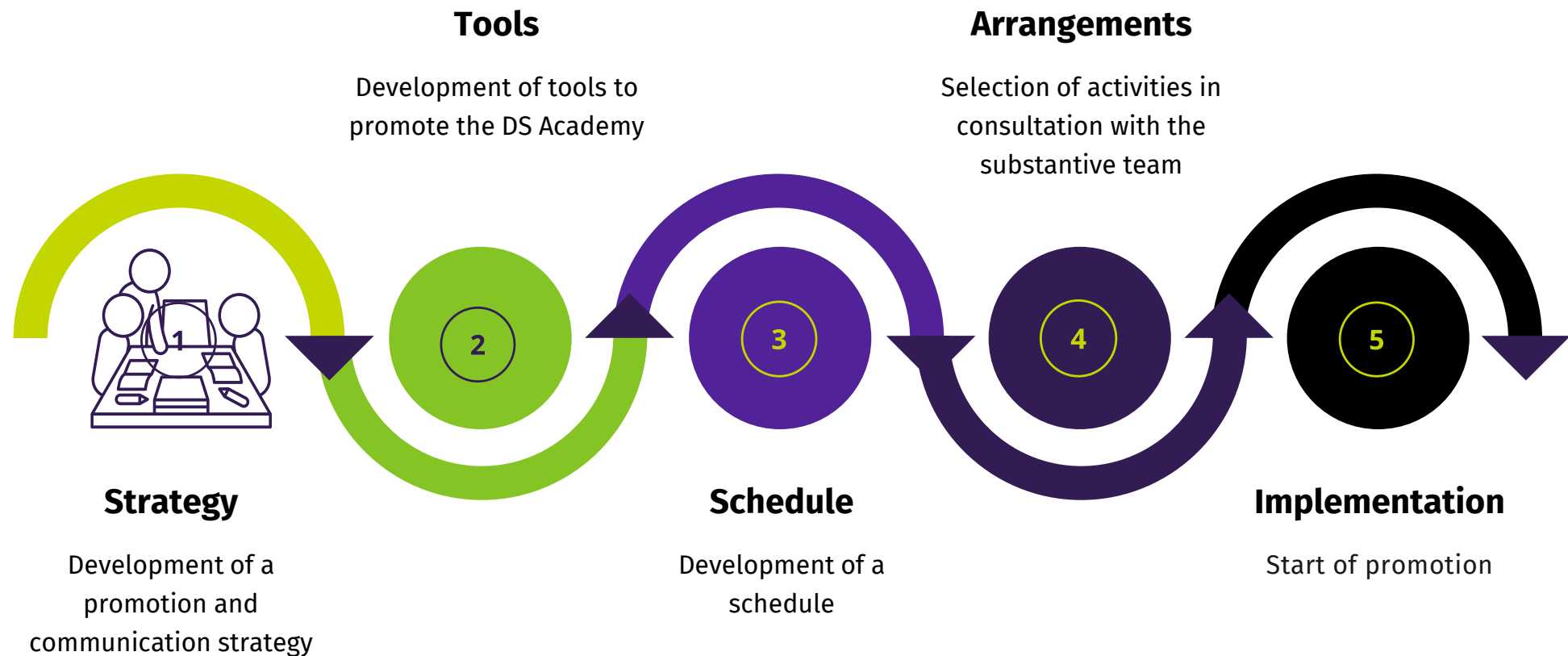
Examples of planned training

- *Ethical and legal issues of data acquisition*
- *Project management*
- *Data security management - introduction to SQL Server Security*
- *Database administration – SQL, system databases, user databases*
- *Relational databases - SQL*
- *Time series modeling*
- *Natural Language Processing - natural language processing, issues of artificial intelligence and linguistics*
- *Privacy Enhancing Techniques – technologies that enhance privacy by minimizing the use of personal data, maximizing data security and empowering individuals*

- *Ethical and legal issues of data acquisition*
- *Project management*
- *Data security management - introduction to SQL Server Security*
- *Database administration – SQL, system databases, user databases*
- *Relational databases - SQL*
- *Time series modeling*
- *Natural Language Processing - natural language processing, issues of artificial intelligence and linguistics*
- *Privacy Enhancing Techniques – technologies that enhance privacy by minimizing the use of personal data, maximizing data security and empowering individuals*



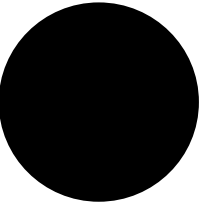

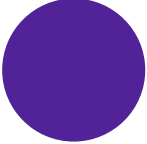



Communication and Promotion Team (1)

From strategy of promotion and communication to implementation



Communication and Promotion Team (2)

Methods of promotion

- Newsletter with mailing 
- Banners 
- Infographics 
- Correspondence 
- Webinars and live 
- Podcasts 
- Promotional videos 
- Information meetings 

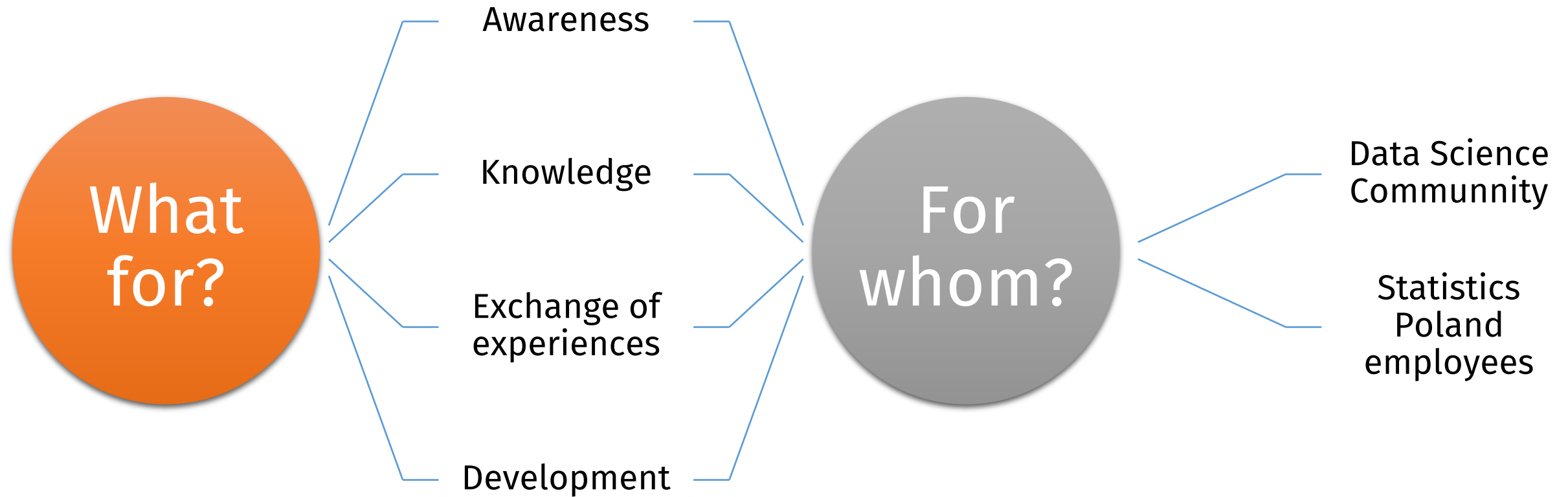
Communication and Promotion Team (3)

Data Science Academy Slogan

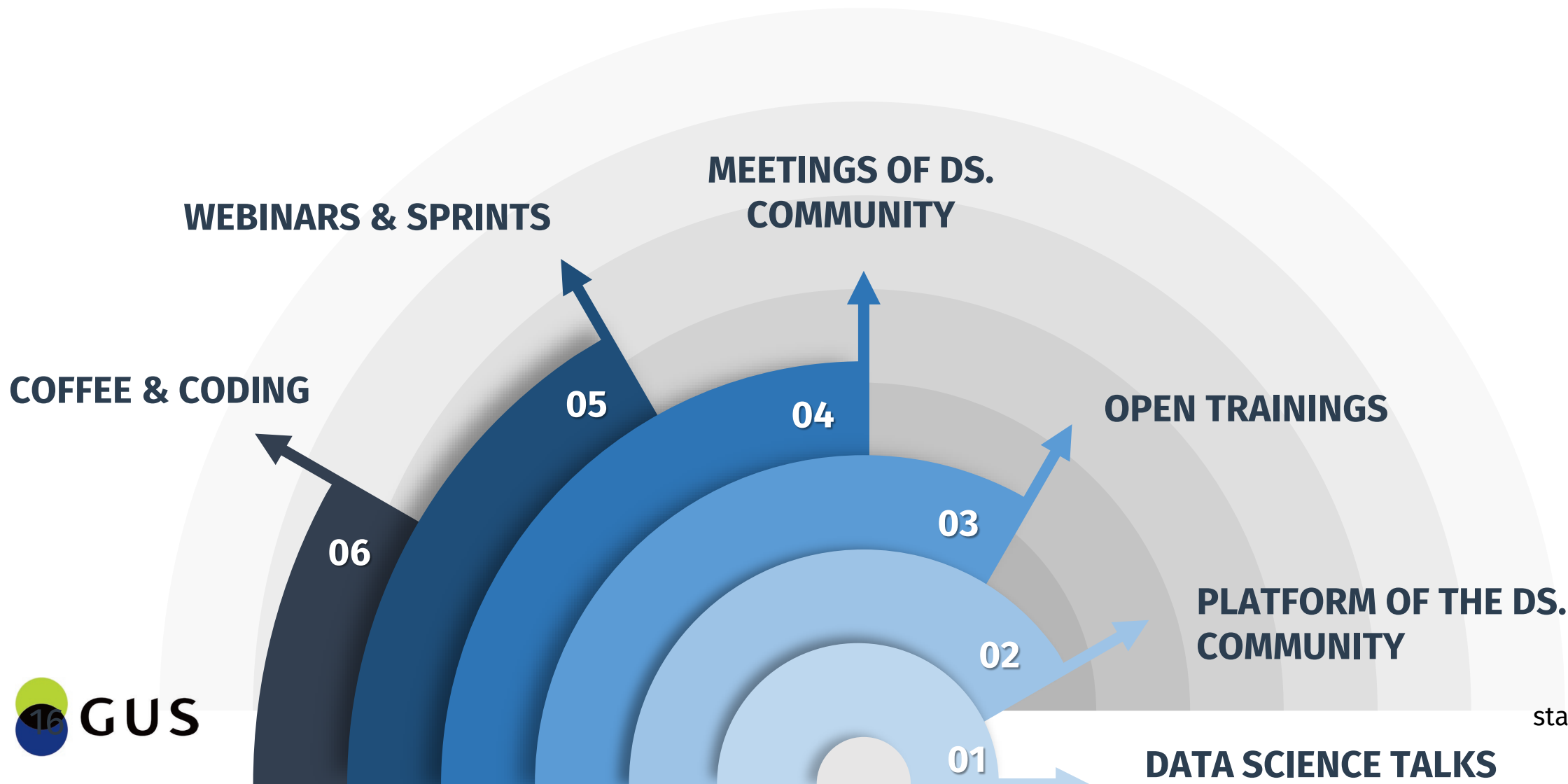
Data Science Academy

Your development in your hands

Building a Data Science community (1)



Building a Data Science community (2)



Action Learning, i.e. learning by doing

- The Action Learning method is an educational process developed by British professor Reginald Revans in the 1940s.
- It involves engaging the team in work on a real, comprehensive and critical problem and taking action, during which there is an intensive learning process of individuals, the team and the organization.
- It is used to develop effective, creative and flexible solutions.
- A characteristic feature of the method is that it enables people with different skills, competences and experience to work together to solve a given problem, develop their leadership skills, contributing to the improvement of team effectiveness.

What has already been done

Completed recruitment
(stage I and II)

"Rules regarding the
remuneration of employees
of public statistics services
units performing tasks for
the Data Science Academy"

Implementation of GitLab to
monitor research work

Standardization of IT
architecture – defined
environment, tools, solution
tests








Series of meetings:
organizational, with
superiors, competence
assessment, and
development meetings

DSA Website
DSA Newsletter

Data Science Talks

Building a DS community

ADS roles and benefits matrix

	Level of expertise	an increase in the basic salary multiplier by 1,25	an increase in the basic salary multiplier by 0,7	conducting training sessions - task allowance PTWstat
Data Scientist	expert			
Data Scientist	specialist			
Data Engineer	expert			
Data Engineer	specialist			
Data Analyst	expert			
Data Analyst	specialist			
Line Manager	expert			

Without DSA

Massive departures of employees with high competences

Delays in the publication of research results

Deteriorating quality of statistical research

Lack of exploration of new research areas

Reducing the number of new projects

Stagnation

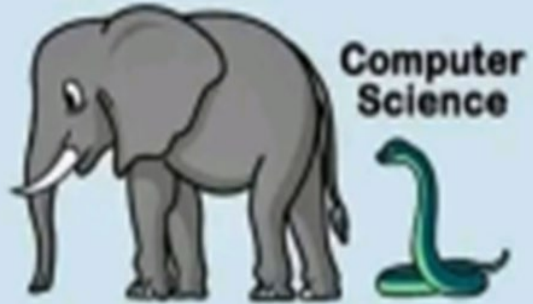
Marsus

Questions for the discussion

1. How to attract talented people to work in statistics?
2. How do you solve trade-off between insourcing and outsourcing of data processing and IT programming?
3. How to motivate staff to use new approaches to data analysis?

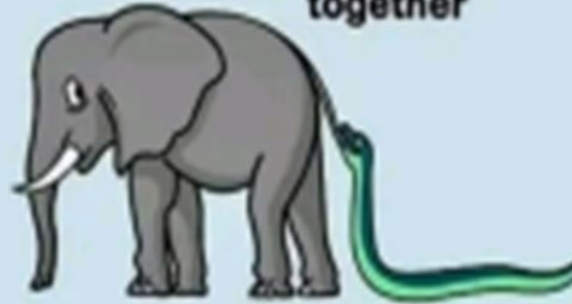


Statistics

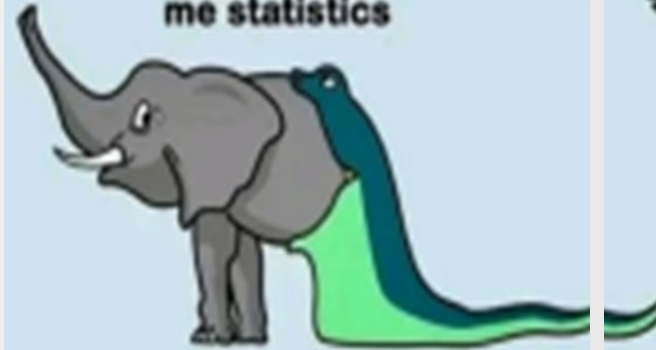


Computer
Science

We will work
together



Please teach
me statistics



Now I m
DATA SCIENTIST



Thank you !

Dominika Rogalińska – d.rogalińska@stat.gov.pl

Anna Borowska – a.borowska@stat.gov.pl